Brain-inspired computer vision with applications to pattern recognition and computer-aided diagnosis of glaucoma

Guo, Jiapan

Chapter 1

Introduction

Probably you have heard about the game named "Spot the differences". In this game, the player needs to locate all differences between two similar pictures, such as, the ones shown in Fig. 1.1. It took me four minutes to locate the ten minutiae differences hidden in these two images. Such a game raises two main insights; the capability (effectiveness) of the visual system of the brain to detect subtle differences in visual scenes and the tediousness (inefficiency) to execute such a task. These considerations have motivated a great deal of literature that uses the brain as a source of inspiration to design computational models of certain behaviour of neurons in visual cortex and the applications of such models to various problems, including assisting domain experts in largely repetitive visual-based tasks. Examples of such applications include machine vision for automatic quality inspection and computer-aided diagnosis in the medical field.

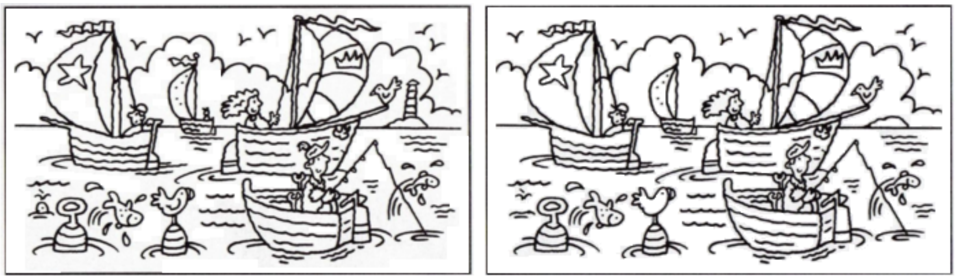


Figure 1.1: An example of the game "Spot the differences" in which the player needs to locate ten differences between two pictures[1].

It is remarkable that more than one third of the brain is devoted to the processing of visual information and that there are more neurons connected to the visual system than the other four senses altogether (Grady, 1993). According to Fixot (1957), when we are awake with eyes open, our vision takes around two thirds of all electrical activity in the brain with 2-3 billion neuron firings per second. Vision starts at the retina with millions of photoreceptors (cones and rods) that are sensitive to

[1]The pictures are taken from: `http://tinyurl.com/n4z39ca`.

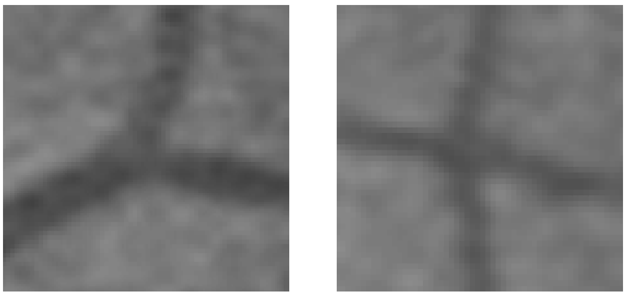
light and color. The photoreceptors transduce light into neural signals that are then transmitted to bipolar and ganglion cells for preprocessing before reaching the lateral geniculate nucleus (LGN). The signals of the LGN cells are then projected to the primary visual cortex, which is the first stage of both the ventral stream (mostly concerned with object recognition) and the dorsal stream (mostly concerned with localization and tracking). Hubel and Wiesel (1962, 1974) were the first who started unveiling the behaviour of neurons in visual cortex. In 1962 they found out by means of single cell recordings in visual cortex of a cat, that there are neurons in this area that are orientation-selective. Their breakthrough motivated lots of research activities of similar nature by other groups who studied the behaviour of neurons in these and other areas of these two streams, and brought insights into other fields which are related to cognition and visual experience (Wade, 2010).

There is psychophysical evidence that vision system is more sensitive to segments of high curvatures than straight segments (Attneave, 1954). Among various discoveries, a phenomenon that struck me most is the inhibition mechanism that is experienced at the neuronal level. Brincat and Connor (2004) discovered the underlying principals of the shape selectivity of neurons in inferotemporal cortex (IT) and that the firings of such neurons can be suppressed by the presence of additional curvatures in their receptive fields. Such an inhibition mechanism is important to the brain, as it facilitates sparseness in the representation of information that may result in an increase in the storage capacity and a higher number of patterns that can be discriminated while keeping the energy consumption to a minimum (Rolls and Treves, 1990). In the first part of this thesis, I focus on developing a computational model that is inspired by the shape-selective neurons in inferotemporal cortex. It turns out that the inhibition property that such shape-selective neurons exhibit improves the discrimination ability, as I demonstrate in applications that require the recognition of complex shapes.

Besides the visual system of the brain, my research interests also include the development of algorithms that can assist domain experts in largely repetitive tasks, in particular ophthalmologic eye screening. Fundus photography is a noninvasive way to inspect the abnormalities associated with eye diseases. It is vital to monitor the progression of certain eye conditions or diseases, such as diabetic retinopathy, age-macular degeneration, and glaucoma, among others. In a population-based screening setting, a huge collection of retinal fundus images is acquired, which requires tedious work on the manual assessment of eye conditions. In the case of glaucoma screening, more than 95% of the people do not have the condition. Thus, an automatic system would be beneficial to rule out all the healthy cases in order for ophthalmologists to concentrate on the final diagnosis and treatment of suspicious glaucomatous subjects. In the second part of this thesis, I worked on the analysis of retinal fundus images based on the extraction of glaucomatous features. My aim is to develop a systematic approach for assisting ophthalmologists in glaucoma screening.

1.1 Scope

Inspired by the function of shape-selective neurons in area V4 of visual cortex, Azzopardi and Petkov (2013b) proposed a computer vision model, which they called COSFIRE (Combination of shifted filter responses). It builds up a hierarchical representation of visual patterns from simple edges/lines to more complex shapes, like combinations of curvatures, that may form whole objects. It has been applied in various computer vision applications for object detection and pattern recognition (Azzopardi and Petkov, 2013b). COSFIRE filters are trainable in that a filter can be configured to be selective for any given pattern, and then that filter will be able to detect the same and similar patterns. While this approach can be advantageous in certain applications, it may be less beneficial in others. For instance, in retinal images, a COSFIRE filter that is selective for a bifurcation/T-junction (Fig. 1.2a) also gives a strong response to a crossover (Fig. 1.2b) of vessels. This indicates that a COSFIRE filter is not robust to distinguish the preferred pattern (e.g T-junction) from other patterns (e.g. crossover) that contain all the contour parts of the preferred pattern.



(a) (b)

Figure 1.2: Example of (a) a bifurcation and (b) a crossover in a retinal fundus images.

The work presented in the first part of this thesis is inspired by the characteristics of shape-selective neurons in inferotemporal cortex. Such neurons respond to complex shapes that are combinations of curvatures in a certain geometrical arrangement. Their responses, however, can be inhibited by the presence of some curvatures in specific locations. Such a phenomenon in the brain occurs in different layers of visual cortex according to neurophysiological evidence (Hubel, 1988;

Brincat and Connor, 2004; Rolls and Treves, 1990; Hubel and Wiesel, 1968; Bolz and Gilbert, 1986). These studies have motivated the work presented in this thesis. In particular, the first research question that I posed is:

1. How can a filter be designed to exhibit properties of inhibition similar to that of the shape-selective neurons in IT cortex?

This question arouse my interest in the design of a novel computational model of shape-selective neurons that also receive inhibitory inputs. We augment the original COSFIRE filter, inspired by neurons that receive only excitatory inputs, with an inhibition mechanism in order to improve its discrimination ability. Such an inhibition-augmented model responds exclusively to patterns that are similar to the one used for configuration. In Chapter 2, we elaborate on the construction of the inhibition-augmented shape-selective model. The configuration of such a model comprises the automatic selection of afferent low-level filters that provide excitatory and inhibitory inputs and the determination of certain parameters that control the trade-off between generalization and selectivity.

After having established an algorithm to configure inhibition-augmented COSFIRE filters, I investigated the following two research questions:

2. In which applications are inhibition-augmented COSFIRE filters most suitable?

3. How would the proposed inhibition-augmented COSFIRE filters perform in comparison to the original COSFIRE filters?

The answers to these two questions are provided in Chapter 2. We demonstrate the effectiveness of the proposed inhibition-augmented COSFIRE filters in three applications, namely the detection of vascular bifurcations in retinal fundus images, the recognition and localization of architectural and electrical symbols and the recognition of handwritten digits. In these applications, it is common to find the contour parts of some objects that are contained within others. For instance, a T-junction is contained within a crossover. In such applications, a COSFIRE filter (without inhibition) that is selective for a T-junction is not sufficiently discriminative since it will respond strongly to T-junctions and crossovers. We also compare the performances of both types of COSFIRE filters, without and with inhibition and show that the proposed inhibition-augmented COSFIRE filters possess better selectivity.

In the second part of the thesis, I focus on a medical image analysis problem, namely the detection of glaucoma from retinal fundus images. Glaucoma is characterized by the degeneration of optic nerve fibres which transmit information from the retina to the brain. It is the second leading cause of blindness, exceeded only by cataract. It, however, brings more severe health issues since the damage of the optic nerve is irreversible. Thus, early diagnosis of glaucoma is crucial in order to enable timely treatment. The manual assessment of retinal fundus images is repetitive and tedious for ophthalmologists to select the suspicious cases since more than 95% of the images contain healthy retinas. In my work, I aim at alleviating the workload of glaucoma experts in the screening procedure and address the following research question:

4. How can computer-vision-based methods assist medical experts with population-based glaucoma screening?

Clinically, the diagnosis and treatment of glaucoma require specialized physicians and sophisticated procedures. However, the initial detection of glaucoma - especially in a polulation-based setting - does not necessarily require all the measurements. Fundus photography is a viable option for population screening of glaucoma and enables the observation of the excavation of the optic nerve head - the hallmark of glaucoma. In our work, we seek to automatically extract interpretable features that are used by ophthalmologists in their daily practice to quantify the changes of the optic nerve head. The following are a couple of research questions that I investigated:

5. What visual features can be used to assess the risk of glaucoma in a computer-aided diagnosis system?

6. How would such a system perform in comparison to medical experts?

These questions are addressed in Chapters 3 and 4. The excavation of the optic nerve head is quantified by the vertical cup-to-disc ratio (VCDR). As illustrated in Fig. 1.3, such a measurement is computed as the ratio $\frac{hc}{hd}$ of the height of the blue ellipse representing the cup and the height of the black ellipse representing the optic disc. The optic disc that is indicated by the black circle, is the two-dimensional view of the optic nerve head, which has an elliptical shape with yellowish color. The cup is the pale, oval region inside the optic disc, which is void of optic nerve fibres. The size of the cup relative to the size of the optic disc gives an indication of the status of the optic nerve head. The VCDR is a commonly used measure to assess the risk of glaucoma (Weinreb and Greve, 2004). Fig. 1.3 shows two retinal fundus im-

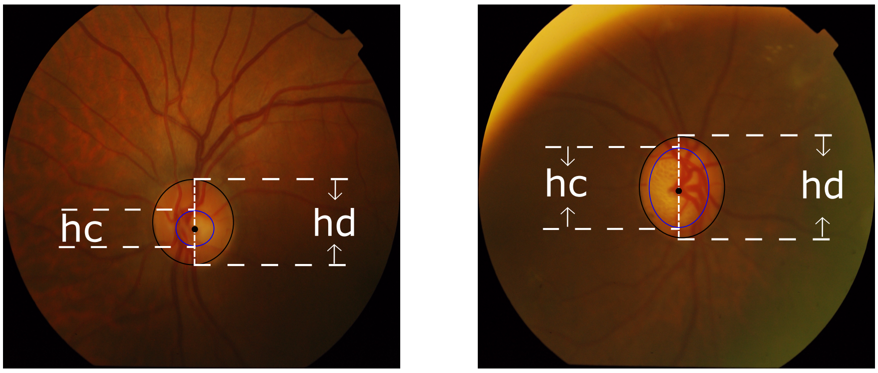


Figure 1.3: Examples of retinal fundus images from a healthy and a suspicious glaucoma retina. The black ellipses indicate the boundaries of the optic discs while the blue ellipses encircle the cup regions. The vertical cup-to-disc ratio is computed as the ratio $\frac{hc}{hd}$ between the height of the cup and that of the disc. The VCDR value is 0.44 and 0.84 for the left and right images, respectively.

ages which are from a healthy retina (VCDR=0.44) and a suspicious glaucoma retina (VCDR=0.84). In Chapter 3 I seek to provide solutions for delineation of the optic disc. Later, in Chapter 4, I focus on the segmentation of the cup and the computation of VCDR values. In order to evaluate the performance of the proposed system, I compare the automatically obtained results on the disc delineation, cup segmentation, and the VCDR with those provided by an experienced ophthalmologist on different data sets.

1.2 Thesis organization

The rest of the thesis is organized in the following way.

Chapter 2 introduces the proposed inhibition-augmented COSFIRE filters for shape detection and object recognition. The responses of shape-selective neurons in the IT area of visual cortex can be inhibited due to the presence of some specific curvatures. Inspired by such neurons, we augment the original COSFIRE filters by adding an inhibition mechanism to improve their discrimination abilities. Finally, we evaluate the performance of the proposed inhibition-augmented COSFIRE filters in three applications: the exclusive detection of vascular bifurcations in retinal fundus images, the recognition and localization of architectural and electrical symbols and the recognition of handwritten digits.

Chapter 3 focuses on an important medical application in ophthalmology. In this part, we seek to apply computer vision techniques to automatically localize the optic disc and estimate its diameter in retinal fundus images. This work is an essential step for computer-aided diagnosis of glaucoma as well as for the detection of other ophthalmologic pathologies, such as diabetic retinopathy, age-related macular degeneration, among others. In this chapter, we investigate two different approaches for the localization and delineation of the optic disc. The first method uses the circular Hough transform to detect candidates for circles and determine the one that encloses the largest proportion of vessels as the approximation of the optic disc. The second approach employs trainable COSFIRE filters which are selective for the divergent points of vessels and disc-like patterns. We then evaluate both methods in several publicly available data sets.

In Chapter 4, we provide a systematic approach on the analysis of retinal fundus images for glaucoma screening based on the vertical cup-to-disc ratio (VCDR). It consists of three steps: optic disc localization and delineation, cup segmentation and computation of the VCDR value. We also provide a reliability score that indicates the confidence by which the system estimates the VCDR in every image. We evaluate the proposed approach on eight test data sets with a total of 1712 images and compare the obtained VCDR values with those manually annotated by an experienced ophthalmologist.

Finally, Chapter 5 summarizes the thesis and provides an outlook of possible directions for future work.